Research Paper

Identification of Host Proteins Predictive of Early Stage *Mycobacterium tuberculosis* Infection

CrossMark

Charles M. Bark [a,c,f,*], Ameur M. Manceur [b], LaShaunda L. Malone [a,c], Mary Nsereko [c], Brenda Okware [c], Harriet K. Mayanja [c,d], Moses L. Joloba [c,e], Isabelle Rajotte [b], Marija Mentinova [b], Phyla Kay [b], Seydina Lo [b], Patrick Tremblay [b], Catherine M. Stein [a,c], W. Henry Boom [a,c], Eustache Paramithiotis [b]

[a] *Tuberculosis Research Unit, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106, USA*
[b] *Caprion Biosciences, 201 President-Kennedy Ave., Montreal, H2X 3Y7, Quebec, Canada*
[c] *Uganda-Case Western Reserve University Research Collaboration, Kampala, Uganda*
[d] *Department of Medicine, College of Health Sciences, Makerere University, Kampala, Uganda*
[e] *Department of Medical Microbiology, College of Health Sciences, Makerere University, Kampala, Uganda*
[f] *Division of Infectious Diseases, MetroHealth Medical Center, Cleveland, OH, USA*

ABSTRACT

The objective of this study was to identify blood-based protein biomarkers of early stage *Mycobacterium tuberculosis* (*Mtb*) infection. We utilized plasma and serum specimens from TB patients and their contacts (age ≥ 12) enrolled in a household contact study in Uganda. In the discovery phase cross-sectional samples from 104 HIV-uninfected persons classified as either active TB, latent *Mtb* infection (LTBI), tuberculin skin test (TST) converters, or persistent TST-negative were analyzed. Two hundred eighty-nine statistically significant (false discovery rate corrected $p < 0.05$) differentially expressed proteins were identified across all comparisons. Proteins associated with cellular immunity and lipid metabolism were induced early after *Mtb* infection. One hundred and fifty-nine proteins were selected for a targeted mass spectrometry assay. A set of longitudinal samples from 52 TST-negative subjects who converted to TST-positive or remained TST-negative were analyzed, and multivariate logistic regression was used to identify unique protein panels able to predict TST conversion with cross-validated AUC > 0.85. Panel performance was confirmed with an independent validation set of longitudinal samples from 16 subjects. These candidate protein biomarkers may allow for the identification of recently *Mtb* infected individuals at highest risk for developing active TB and most likely to benefit from preventive therapy.

1. Introduction

Approximately one-third of the world's population is latently infected with *Mycobacterium tuberculosis*, meaning they do not have symptoms, chest radiographic abnormalities, or other findings of active tuberculosis (TB). People with latent *Mtb* infection (LTBI) are the primary source of future TB cases, and their identification is important for TB control. Diagnosis of LTBI is based on immunological activity suggesting current or previous infection, commonly measured by either the tuberculin skin test (TST) or interferon gamma release assay (IGRA). Neither test is able to differentiate between LTBI and active TB, nor distinguish recent from remote infection. This is an important distinction since recent infection is a strong risk factor for progression to active TB, and in some high incidence areas, the majority of TB cases are likely due to

recent infections from ongoing TB transmission (Chin et al., 1998; Verver et al., 2004). Developing a diagnostic assay that identifies recent *Mtb* infection would allow for targeted treatment of those persons most likely to progress to active TB and is a priority among international TB agencies (Pai and Schito, 2015).

Mass spectrometry (MS) coupled with multiple reaction monitoring (MRM-MS) allows for rapid detection and quantification of proteins with high sensitivity and precision (Hunter and Paramithiotis, 2010). Previous studies have used MRM-MS proteomic assays to identify new biomarkers of LTBI by detecting both *Mtb* (Kruh-Garcia et al., 2014) and human host proteins in peripheral blood (Sandhu et al., 2012; Zhang et al., 2014). In these previous cross-sectional studies, LTBI was diagnosed by TST or IGRA, but it was not known when the subject was infected with *Mtb* (recent vs remote infection). In this study, we analyzed blood samples from a prospective TB household contact cohort using MRM-MS to assess the host-protein proteomic profiles in blood from household contacts who converted from TST-negative to TST-positive. Here we report changes in circulating host-proteins as a person develops LTBI.

* Corresponding author at: Tuberculosis Research Unit, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106, USA.
  E-mail address: cmb148@case.edu (C.M. Bark).

2. Materials and Methods

2.1. Study Design and Subjects

Participants for this study were enrolled between 2002 and 2012 in the Kawempe Community Health Study (KCHS), a prospective cohort of adult pulmonary TB index cases and their household contacts (Whalen et al., 2006). Kawempe is located in Kampala, Uganda, which is a high-burden country with an annual TB incidence of approximately 202 TB cases per 100,000 population (World Health Organization, 2016). Index cases were defined as adults (age 18 years and older) with initial episodes of newly diagnosed culture-positive pulmonary TB (Stein et al., 2013). A household contact was defined as a person living within the same building as an index case for at least one week during the three-month period immediately preceding the diagnosis of TB in the index case. Contacts included in this sub-study were age 12 years and older. TST was performed by Mantoux method with 5 TU of purified protein derivative (PPD). After the initial evaluation, participants were evaluated at 1, 3, 6, 12, and 24 months for active TB and with repeat TST at months 3, 6, 12, and 24 if their first and subsequent TST remained negative. Converters were defined as household contacts, with an initial TST < 10 mm at baseline visit, who subsequently converted their skin test to positive (TST ≥ 10 mm and an increment of 6 mm) during follow-up testing. Subjects that remained TST-negative and did not convert their TST were considered to be persistently not infected (NI) (Ma et al., 2014). All subjects with a positive TST (at baseline or conversion during follow-up) were offered treatment with isoniazid preventive therapy (IPT) (10–20 mg/kg or a maximum dose of 300 mg/day) for 9 months. All subjects included in this sub-study were HIV negative. The study was approved by the responsible institutional review boards in Uganda and the U.S and all participants provided written informed consent.

2.2. Sample Processing

Sera and plasma were aliquoted and stored at −80 °C until testing. Samples were grouped into blocks containing each of the clinical groups (active TB, LTBI, converters, and NI) and the order of the samples within each block was randomized. Samples were depleted of abundant proteins using an HSA/IgG column (Agilent Technologies, Mississauga, ON) in tandem with an IgY14 and Supermix (Sigma, Oakville, Ontario) column and the flow through was digested with trypsin (Promega, Madison, WI) at a trypsin to protein ratio of 1:10. The digested samples were freeze-dried, resolubilized, and treated with TCEP [tris(2-carboxyethyl)phosphine] to reduce disulfide bonds. Samples were desalted by solid phase extraction using a 3M Empore C18 desalting plate and distributed into 96-well plates and vacuum evaporated. Peptides were stored at −20 °C until use. Sample processing variability was measured using aliquots of pooled plasma that were inserted at regular intervals among the study samples and taken through the entire analysis. For the Discovery study the process quality control sample CV was 20.45%, and for Verification study 1 and 2 the CVs were 10.99% and 4.3%, respectively.

2.3. Tandem Mass Spectrometry Analysis

Samples were resuspended in 92.5/7.5 water/ACN + 0.2% formic acid and analyzed by LC-MS (nanoAcquity UPLC and Q Exactive mass spectrometer). Peptide separation was achieved using a Waters nanoAcquity Symmetry UPLC Trap column (180 μm × 20 mm, 5 μm particle size) and a Waters nanoAcquity UPLC BEH130 analytical column (150 μm × 100 mm, 1.7 μm particle size). The mobile phases were (A) 0.2% formic acid in water and (B) 0.2% formic acid in acetonitrile. For each sample approximately 2.5 μg was loaded onto the analytical column. Peptides were separated using a 52.5 min gradient (92.5% A to 40% A) at a flow rate of 1.8 μL/min. MS spectra were acquired from 400 to 1800 Da. The MS method consisted of a Full MS scan followed by a dd-MS2 scan of the top 12 ions. The Full MS scan was achieved at a resolution of 70,000 with an AGC value of 3e6 and a maximum IT of 30 ms. The dd-MS2 scan was performed at a resolution of 17,500 with an AGC target value of 5e4 and a maximum IT of 60 ms.

Components, defined as ions that have been identified by their mass/charge ratios, retention times, and intensities but do not yet have sequence information, were detected and matched across all samples using the Elucidator software version 3.3 (Rosetta Biosoftware, Seattle, WA). Presence of unusual modifications was determined using Preview Software (Protein Metrics). Fragmentation patterns were matched to the corresponding sequences found in a combined database containing Human (reviewed entries only) downloaded on 20,121,206 from Uniprot, using Mascot software version 2.2.06 (Matrix Science, Boston, MA). The following parameters were used for the searches: enzyme = trypsin, allowed missed cleavages = 2, peptide tolerance = 20 ppm, MS/MS tolerance = 0.4 Da, variable modifications = Deamidation (N, Q), Oxidation (M), Acetaldehyde (N-term). A decoy reverse database was used to evaluate false positive error rate. Peptide/Protein Teller helped derive the simplest list of proteins to explain observed peptides, at a FDR of 5%. All intensities were log (base e) transformed. 42,654 peptides, whose intensities reached LOD (= 50,000) in at least 15 samples, were included in subsequent analyses. The overall intensity bias was corrected so that, between each pair of samples, the median of the difference of the peptide intensities (in log scale) of the two samples over all the peptides was 0, or close to. For batch-effect correction, a one-way ANOVA model $I_{ij} = M + D_i + \varepsilon_{ij}$ (I: intensity, M: overall interception, and D: batch–factor) was solved and parameters D_i ($i = 1…5$) under the constraint of $\Sigma_{i=1}^{10} (N_i * D_i) = 0$ were obtained; the D_i's were then subtracted from the normalized intensities to form the 'batch-effect corrected' intensities. Intensities below an intensity threshold (IT), IT = 50,000, were transformed to avoid spurious large fold changes as follows: $I \rightarrow (I + IT) / 2$ (i.e. intensity in the range of (0, IT) was linearly mapped to the range of (IT/2, IT)). ANOVA analysis was then applied to identify differentially expressed peptides. ANOVA model: $I_{ij} = M + T_i + \varepsilon_{ij}$ where I is peptide intensity, M is overall average intensity, t is 'condition' factor, and ε random error (Keeping, 1995; Montgomery, 2001). False detection rate and q-value were calculated based on the p-values obtained from the ANOVA using Storey's method to make multiple testing adjustments (implemented in MATLAB) (Storey, 2002). 'Post hoc' contrast analyses were conducted using Tukey's hsd method to calculate p-values for each pair wise comparison (Hochberg and Tamhane, 2008). A protein level analysis was then applied by introducing a 'peptide factor' in the ANOVA used above: $I_{ijk} = M + T_i + P_j + \varepsilon_{ijk}$ where I is protein intensity, M overall constant, T 'condition', and P peptide factor. The number of the levels for P is protein-dependent, equal to the number of peptides for the protein. Proteins were considered to be differentially expressed if they met the following thresholds: p- and q-values

Table 1
Characteristics of included subjects by study phase and clinical cohort.

	Discovery (N = 104)				Verification 1 (N = 52)		Verification 2 (N = 16)	
	Non-Infected	Converter	LTBI	Active TB	Non-infected (longitudinal)	Converter (longitudinal)	Non-infected (longitudinal)	Converter (longitudinal)
n	8	21	38	37	15	37	3	13
Age (median)	20	25	25	25	20	16	25	21
Sex (% male)	63	24	40	43	40	27	0	31
BMI (median)	21	22	22	19	21	21	21	22

<0.05 and differential intensity (DI) > 1.1-fold change (or <0.9 for decrease).

2.4. Multiple Reaction Monitoring Mass Spectrometry

The MRM experiments were done as described previously (Achkar et al., 2015). Three hundred ninety-two (392) peptides representing 163 host proteins were synthesized (JPT Peptide Technologies, Berlin, Germany) and resolubilized in 25/75 DMSO/water (v/v), pooled and diluted with water + 0.2% formic acid to a concentration of 200 pmol/mL. This peptide mix was used to develop the MRM assay on a nanoAcquity UPLC (Waters) coupled to a QTRAP 5500 mass spectrometer (AB Sciex). Peptide separation was achieved using a BioBasic C18 column (Thermo) (320 μm × 150 mm, 5 μm particle size). The mobile phases were (A) 0.2% formic acid in water and (B) 0.2% formic acid in acetonitrile. Gradient time was 30 min and the flow rate was 10 μL/min. The optimal 2

Panel 1

Category	Protein	CO v NI	LTBI v NI	ATB v NI	ATB v LTBI
Acute phase response	HP	1.26	1.34	6.26	4.67
	ORM1	1.25	1.37	3.83	2.80
	SAA4	1.28	1.46	1.29	0.88
	C1S	1.14	1.20	1.18	0.98
	C1R	1.16	1.23	1.13	0.92
	F5	1.15	1.14	1.01	0.89
	SAA1	1.39	1.93	13.16	6.83
	SERPINF1	1.03	1.12	0.88	0.78
	SAA2	0.93	1.09	3.15	2.89
	APCS	1.07	1.20	2.33	1.93
	ORM2	1.19	1.28	1.95	1.52
	AMBP	1.05	1.05	1.52	1.45
	F9	0.91	1.10	1.24	1.13
	AHSG	0.94	1.03	0.54	0.53
	TF	1.04	1.00	0.49	0.94
	IL1RAP	1.01	0.98	0.44	0.44
	CD163	1.13	1.08	0.97	0.90
Inflammatory response	S100A8	1.26	1.29	2.64	2.04
	S100A9	1.22	1.28	3.51	2.74
	CD5L	0.89	0.83	0.84	1.01
	SOD3	0.85	0.73	0.59	0.81
	CAT	0.79	0.60	0.40	0.67
	CPN1	1.15	1.15	1.07	0.93
	ANPEP	0.88	0.82	0.68	0.83
	ATRN	0.92	0.94	0.63	0.68
	ALCAM	1.00	0.74	0.41	0.55
	KRT14	0.66	0.53	1.08	2.03
Complement	CRP	1.39	1.73	8.73	5.05
	C2	1.12	1.16	1.21	1.04
	C3	1.26	1.34	1.20	0.89
	CLU	1.13	1.19	0.92	0.77
	C7	0.85	0.84	0.78	0.92
	MBL2	0.70	0.85	0.64	0.75
	C9	1.18	1.44	2.79	1.93
	FCN3	1.09	1.38	1.49	1.08
	KNG1	1.28	1.32	0.72	0.54
	MASP1	0.96	1.00	0.79	0.79
	CFD	0.89	0.92	0.51	0.55
	CR2	1.04	0.73	0.48	0.66
	C4BPA	1.10	1.15	2.05	1.78
	MASP2	1.10	1.28	0.98	0.77
Coagulation	SERPIND1	1.19	1.20	1.88	1.57
	VWF	1.34	1.08	1.06	0.98
	PLEK	1.46	1.23	0.92	0.75
	FLNA	1.29	1.18	0.68	0.58
	PF4	0.72	0.74	0.62	0.84
	F11	1.09	1.22	1.09	0.89
	F12	0.93	0.82	0.84	1.02
	PTGDS	0.89	0.92	0.80	0.87
	GP1BA	1.02	1.01	0.77	0.76
	MMRN1	0.80	0.79	0.74	0.93
	F7	1.17	1.12	0.68	0.60
	ADAMTS13	0.85	0.81	0.55	0.68
	GP5	0.91	0.94	0.54	0.58
	LRP1	1.12	0.94	0.43	0.46
	SERPINA5	1.02	0.83	0.41	0.50
	FGL2	0.50	0.52	0.40	0.77
	PROCR	0.87	0.84	0.31	0.37
	FGA	1.14	1.04	1.23	1.18
Oxidative stress	S100A12	1.85	1.98	6.31	3.18
	PRDX2	0.73	0.55	0.40	0.72
	PON1	1.20	1.43	0.73	0.51
	AGT	0.97	1.07	0.83	0.78
	SH3BGRL3	1.02	0.97	0.73	0.76
	SEPP1	0.83	0.90	0.71	0.78
	SOD1	0.96	0.69	0.57	0.83
	MMP2	0.94	0.74	0.29	0.39
Innate immunity	LGALS3BP	1.31	1.33	1.65	1.23
	LILRA3	1.62	1.47	1.25	0.85
	VNN1	1.25	1.49	1.18	0.79
	LBP	1.12	1.40	2.44	1.75
	PGLYRP2	1.23	1.24	0.68	0.55
	HPR	1.01	1.06	1.78	1.69
	CSF1R	1.07	1.00	0.57	0.57
	SELENBP1	0.94	0.58	0.32	0.55

Panel 2

Category	Protein	CO v NI	LTBI v NI	ATB v NI	ATB v LTBI
Adaptive immunity	SERPINA1	1.19	1.27	3.32	2.62
	LCP1	1.30	1.30	1.61	1.24
	IGLL5	0.88	0.91	1.29	1.41
	DBH	1.52	1.12	0.92	0.82
	A2M	0.91	1.04	0.81	0.79
	SERPINC1	1.70	1.93	1.21	0.62
	LPA	1.28	1.77	1.14	0.65
	PIGR	0.92	0.77	1.08	1.40
	FCGR3A	1.12	1.20	1.45	1.21
	CD14	1.03	1.09	1.42	1.31
	MMP9	1.11	1.05	1.37	1.31
	ICAM1	1.06	1.17	1.37	1.17
	CPB2	0.98	1.09	1.19	1.09
	PROS1	0.90	1.02	0.90	0.88
	THBS1	0.98	1.00	0.85	0.85
	ICOSLG	0.91	1.00	0.75	0.75
	VCAM1	1.02	1.01	0.68	0.67
	PTPRJ	0.98	0.90	0.58	0.64
	PNP	0.73	0.53	0.46	0.87
	BLVRB	0.98	0.66	0.44	0.67
	THBS4	1.03	0.79	0.31	0.40
	DPP4	0.76	0.68	0.27	0.40
	SECTM1	1.26	1.38	0.22	0.16
	TAGLN2	1.34	1.33	1.04	0.78
Lipid metabolism	APOA4	0.72	0.80	0.57	0.71
	APOB	1.38	1.41	0.96	0.68
	APOE	1.15	1.29	1.00	0.77
	LCAT	1.17	1.20	0.91	0.76
	APOF	1.30	1.31	0.99	0.76
	APOD	1.13	1.31	0.81	0.62
	APOA1	1.03	1.11	0.71	0.64
	APOC1	0.97	1.09	0.72	0.66
	PCSK9	0.99	0.77	0.65	0.85
	APOC2	0.94	1.04	0.63	0.61
	APOC4	0.85	0.80	0.55	0.69
	APOA2	0.97	1.04	0.51	0.49
	CETP	1.03	1.10	1.43	1.31
	PRG4	1.22	1.12	2.09	1.87
Glycolysis	TPI1	0.97	0.79	0.71	0.90
	ALDOA	0.81	0.83	0.65	0.78
	LDHB	0.91	0.79	0.53	0.67
	LDHA	0.90	0.62	0.49	0.79
Carb.	CTBS	1.06	1.03	0.63	0.61
	NAGLU	0.90	0.92	0.56	0.61
Tissue development & remodelling	HRNR	0.30	0.24	0.24	0.99
	KRT9	0.72	0.65	0.62	0.96
	FCGBP	0.80	0.87	0.85	0.97
	ECM1	0.84	0.81	0.92	1.13
	LUM	0.88	0.90	0.57	0.63
	LRG1	1.18	1.34	2.90	2.16
	KRT5	0.72	0.68	0.78	1.16
	CA2	0.83	0.61	0.39	0.63
	S100A4	1.20	0.77	0.16	0.21
	COL1A1	0.66	0.49	0.17	0.34
	CD248	0.90	0.74	0.19	0.26
	RTN4RL2	0.77	0.76	0.35	0.46
	COL6A1	0.76	0.77	0.36	0.47
	PTPRG	1.01	1.05	0.40	0.38
	KRT2	0.61	0.66	0.41	0.62
	MEGF8	0.85	0.78	0.41	0.52
	COLEC11	1.25	0.97	0.42	0.44
	EFEMP1	1.13	1.12	0.45	0.40
	NCAM2	0.80	0.86	0.48	0.56
	SEMA4B	0.96	1.03	0.50	0.48
	OGN	1.06	0.96	0.51	0.53
	COL6A3	1.01	0.90	0.52	0.57
	ABI3BP	0.79	0.75	0.53	0.71
	SPP2	1.05	1.05	0.53	0.50
	DAG1	0.81	0.87	0.54	0.62
	COMP	1.11	1.00	0.55	0.56
	SPARCL1	1.10	1.06	0.56	0.53
	CRTAC1	1.04	1.01	0.62	0.61
	CD109	1.27	1.08	0.65	0.60
	BTD	0.99	1.07	0.65	0.61
	VASN	0.87	0.87	0.73	0.84
	QSOX1	1.06	1.13	0.82	0.72
	NRP1	0.97	1.09	1.22	1.12
	PLS1	1.11	1.08	1.45	1.35
	FAM20C	1.22	1.41	1.81	1.29
	PFN1	1.33	1.09	0.74	0.68

Panel 3

Category	Protein	CO v NI	LTBI v NI	ATB v NI	ATB v LTBI
Cell proliferation	IGFBP3	0.81	0.84	0.41	0.49
	IGFBP5	0.43	0.46	0.27	0.60
	IGF1	0.70	0.67	0.42	0.62
	IGFBP6	0.95	1.01	0.74	0.73
	SPARC	0.97	0.98	0.73	0.75
	BCHE	1.04	1.03	0.72	0.71
	CHL1	1.06	1.01	0.64	0.63
	FSTL1	0.86	0.99	0.55	0.56
	IGF2	1.05	1.01	0.53	0.52
	ENO1	1.07	0.98	0.41	0.42
	PDGFRB	0.61	0.69	0.39	0.56
	AIFM1	1.17	1.13	0.26	0.23
	LTBP1	0.79	0.64	0.20	0.32
	AK1	0.79	0.47	0.10	0.21
	IGFBP2	0.90	1.13	1.55	1.38
Cell migration & adhesion	TNXB	0.85	0.83	0.40	0.49
	GPLD1	0.79	0.98	0.58	0.59
	FBLN1	0.85	0.81	0.53	0.65
	IGFALS	1.00	0.99	0.77	0.78
	VCL	1.12	1.06	0.73	0.69
	PVR	0.72	0.84	0.71	0.85
	AOC3	0.92	0.91	0.65	0.72
	NEO1	0.99	1.06	0.63	0.59
	ANGPTL3	1.44	1.26	0.55	0.44
	TLN1	1.22	1.08	0.55	0.51
	GSN	0.97	0.98	0.54	0.55
	NCAM1	0.99	0.95	0.53	0.56
	CDH5	1.13	1.08	0.52	0.48
	HSPG2	1.06	0.85	0.51	0.60
	TPM3	1.20	1.12	0.50	0.45
	CDH13	1.15	0.99	0.49	0.49
	DSG2	1.17	1.19	0.48	0.40
	TPM4	1.26	1.16	0.47	0.41
	MCAM	1.14	1.13	0.43	0.38
	POSTN	0.98	0.81	0.38	0.48
	CNTN1	0.94	0.84	0.38	0.45
	OMD	0.79	0.67	0.30	0.45
	ADAMDEC1	0.87	0.80	0.29	0.36
	ACAN	0.56	0.52	0.24	0.45
	CLSTN1	0.92	0.69	0.15	0.22
	MYOC	0.74	0.50	0.13	0.26
	PSG2	1.21	1.21	1.98	1.64
	VTN	1.10	1.16	0.75	0.65
Glycosylation	MAN1A1	1.37	1.40	1.85	1.33
	FUCA1	0.91	0.91	0.63	0.69
	MAN2A1	1.01	0.96	0.54	0.57
	FUCA2	0.73	0.84	0.50	0.59
	B3GNT2	1.10	0.92	0.31	0.34
Proteolysis	PZP	1.49	1.28	1.83	1.43
	CNDP1	1.24	1.21	0.87	0.72
	ITIH3	1.22	1.30	1.77	1.36
	SERPINA6	1.02	1.13	1.07	0.94
	PI16	0.83	0.73	0.30	0.42
	CPQ	0.98	0.63	0.18	0.28
	UBE3B	0.80	0.83	0.33	0.40
	PCOLCE	0.97	0.89	0.42	0.47
	XPNPEP2	0.76	0.67	0.42	0.63
	CST6	0.84	0.86	0.45	0.52
	TIMP2	0.66	0.62	0.50	0.80
	SERPINA4	0.95	1.00	0.54	0.54
	HGFAC	0.98	0.93	0.58	0.63
	PEPD	1.08	1.07	0.67	0.62
	CST3	0.96	1.07	0.78	0.73
	FETUB	1.04	1.09	1.32	1.21
	GGH	1.15	1.14	1.44	1.26
	CPN2	1.10	1.02	0.92	0.90
Transport	TTR	2.86	2.51	1.67	0.67
	SHBG	1.54	1.49	1.13	0.76
	ALB	0.78	0.89	0.38	0.43
	AFM	0.87	0.85	0.45	0.53
	HBB	0.98	0.66	1.21	1.83
	HBA1	0.87	0.55	0.97	1.76
	SNCA	0.89	0.58	0.21	0.36
	ZC3H3	0.44	0.42	0.30	0.71
	LAMP1	0.92	0.96	0.47	0.49
	LYVE1	0.93	0.99	0.52	0.53
	CLEC3B	1.00	1.06	0.55	0.52
	CACNA2D1	1.15	0.96	0.58	0.60
	HSPA5	0.97	0.99	0.72	0.72
	TFRC	1.02	0.96	0.76	0.79
	FTL	0.94	1.04	2.39	2.31
	PDLIM1	1.24	1.14	0.75	0.66
Other	CA1	0.92	0.60	0.35	0.58
	RNASE1	1.06	0.97	0.48	0.49
	MST1	1.07	1.04	0.91	0.87

Fig. 1. Cross-sectional comparison of changes in sera from non-infected (NI), baseline TST-negative future converters (CO), baseline TST-positive (LTBI), and active TB (ATB) groups. Shown are protein expression change ratios. The significant changes (p < 0.05) are color-coded red for increased and blue for decreased expression ratios, respectively.

transitions per peptide were determined using selected reaction monitoring (SRM)-triggered MS/MS on a QTRAP 5500 instrument (AB Sciex). The MRM-MS/MS method was developed by calculating, for each peptide, the precursor mass of the doubly and triply charged peptide ions and the first and second y fragment ions with a m/z greater than [m/z (precursor) + 20 Da]. If these transitions were observed during the MRM scan a full MS/MS spectrum of the precursor peptide ion was acquired. The two most intense fragment ions (b or y fragment ions only) in the MS/MS spectrum and its elution time were determined for each acquired peptide and the collision energy (CE) was optimized for all of the chosen transitions. Sample analysis utilized an MRM detection window of 140 s ($\pm$70s around the RT of each peptide) and a cycle time of 1.5 s. The developed MRM assay was then applied for the analysis of the study samples.

Expression analysis of MRM-MS data acquired from the verification samples was performed using R version 2.14.0, platform x86_64-pc-mingw32/x64 (64-bit). An IT below which the measure is deemed less reliable was determined empirically and set to 10,000 pre-normalization. A detection rate (DR) was defined as the proportion of samples within a group with a raw intensity value greater or equal to the IT. Transitions with DR below 50% for one of the two groups being compared were excluded from expression analysis. Differential intensity (DI) ratios were calculated in pair wise comparisons for each transition as the ratio of the average normalized intensities of each group. Prior to calculating the DI ratios, all intensity values that were below IT in the raw data prior to normalization were replaced by the half-IT value. Student's t-test was applied for the expression analysis. Protein-level statistics were also computed by linearly combining the transitions of a given protein into a single variable and then applying a t-test. The average was a weighted average, where the weights were derived from the first principal component of the transition covariance matrix. The calculation of q-values was done using function "qvalue" from Storey's package "qvalue" version 1.24.0 (Storey, 2002).

2.5. Network Analysis

Data were analyzed using Ingenuity Pathway Analysis (IPA; Ingenuity® Systems, Redwood City, CA). Right-tailed Fisher's exact test was used to calculate a p-value determining the probability that each biological function and/or disease assigned to that data set was due to chance alone.

	Protein	Becomes TST+	Remains TST-
Acute phase	ORM1	1.63	1.95
	CD163	1.19	1.34
Inflammatory response	SOD3	1.29	1.78
	CPN1	1.23	1.16
	ATRN	1.20	1.17
	S100A8	1.13	1.75
	S100A9	1.12	1.78
Complement	C7	1.37	1.40
	CR2	0.93	2.41
	FCN3	1.00	1.64
	MBL2	1.17	1.22
Coagulation	VWF	1.42	1.35
	HABP2	1.31	1.23
	FLNA	1.33	1.01
	SERPINA5	1.26	1.12
	PROCR	1.23	1.14
	PLEK	1.13	0.86
Oxidative stress	GPX3	1.31	1.29
	HYOU1	1.23	1.20
	PON1	1.28	1.00
	SEPP1	1.25	1.19
	S100A12	1.03	1.74
Innate immunity	LBP	1.46	1.25
	PGLYRP2	1.38	1.21
	VNN1	1.25	1.20
Adaptive immunity	THBS4	1.39	1.20
	CD14	1.28	1.27
	PROS1	1.27	1.22
	LCP1	1.23	1.31
	VCAM1	1.23	1.29
	SECTM1	1.30	1.19
	LPA	1.21	1.08
	CPB2	1.20	1.18
	DBH	1.02	2.00
Lipid metabolism	APOA4	1.76	1.37
	APOE	1.76	1.32
	APOA1	1.70	1.40
	PCSK9	1.42	2.59
	PRG4	1.29	1.24
	APOD	1.27	1.04
	APOC3	1.22	0.97
	LCAT	1.21	1.16
Glycolysis	LDHA	1.31	1.12
Transport	CLEC3B	1.29	1.22
	PTGDS	1.24	1.22
	CACNA2D1	1.21	1.00
	LAMP1	1.19	1.20
	HBA1	0.79	1.01

	Protein	Becomes TST+	Remains TST-
Tissue development & remodelling	MMP2	1.47	1.20
	ECM1	1.43	1.34
	LRG1	1.28	1.34
	LUM	1.25	1.22
	BTD	1.23	1.26
	COLEC11	1.23	1.20
	COMP	1.35	1.19
	FCGBP	1.35	1.01
	CD248	1.34	1.18
	CNTN1	1.32	1.01
	COL1A1	1.25	1.17
	MINPP1	1.24	1.19
	PTPRG	1.22	1.15
	MEGF8	1.21	1.08
	VASN	1.21	1.19
Cell proliferation	IGF2	1.58	1.70
	ENO1	1.32	1.03
	IGFBP3	1.24	1.16
	BCHE	1.21	1.14
	AK1	1.08	2.42
Cell migration & adhesion	ZYX	1.64	1.72
	POSTN	1.40	1.32
	MYOC	1.30	1.26
	TNXB	1.30	1.23
	TGFBI	1.25	1.23
	ADAMDEC1	1.32	1.15
	CDH13	1.24	1.19
	IGFALS	1.24	1.16
	OMD	1.23	1.10
	DSG2	1.22	1.16
	NID1	1.21	1.19
	GPLD1	1.20	1.13
	MCAM	1.20	1.14
	CLSTN1	1.15	1.48
	SELL	1.15	1.25
	CDH5	1.18	1.23
Glycosylation	FUCA2	1.29	1.19
	MAN1A1	1.19	1.23
Proteolysis	PCOLCE	1.30	1.26
	CPN2	1.30	1.20
	CST6	1.25	1.19
	CNDP1	1.22	1.19
	CPQ	1.22	1.05
	PEPD	1.22	1.17
Other	MST1	1.20	1.23
	CKM	1.65	1.19
	YWHAE	0.34	4.07

Fig. 2. Longitudinal comparison of changes in selected* plasma proteins from baseline TST-negative subjects ($N = 52$) that converted to TST-positive ($N = 37$) or remained TST-negative ($N = 15$). Shown are protein expression change ratios between each individual's baseline TST-negative sample and the TST-positive conversion sample or corresponding TST-negative sample. The significant changes ($p < 0.05$) are color-coded red for increased and blue for decreased expression ratios, respectively. *We selected 159 proteins from the cross-sectional discovery phase using the following combination of biological and statistical criteria. All of the significantly differentially expressed proteins from the baseline converter vs NI and LTBI vs NI comparisons were selected along with the most differentially expressed proteins from comparisons to active TB. We also included the significant proteins described in the MRM-MS assay of Achkar et al. (2015).

2.5.1. Panel Selection and Receiver Operating Characteristics

The area under the curve (AUC) was the result of a nested cross-validation procedure that used stratified sampling to split the data at random into five test sets. For each test set, the remaining four fifths (4/5) of the data were defined to be that test set's corresponding training set. To facilitate the selection of panels of interest, the following procedure was carried out for all possible panels with up to four proteins out of all the candidate biomarkers. Each training set was again split at random by stratified sampling into two halves. One half was used to fit a logistic regression model, which was then used to calculate out-of-sample predictive scores for the other half. This random half-and-half split procedure was repeated a number of times equal to three times the sample size of the training set; out-of-sample predictive scores and the corresponding true outcomes were aggregated over all random splits and AUCs were estimated from these. Since there are five training sets, five such AUC estimates were generated for each panel, which were then averaged. Panel selection was carried out by examining various summaries of protein performance, and also direct examination of the panels with the best AUC estimates. We used a predetermined selection parameter of AUC > 0.85. To compute the final AUC estimates of the selected panels, each test set was scored by a logistic regression model fit to the corresponding training set; the resulting out-of-sample predictive scores and true outcomes aggregated over all five test sets, forming the final set from which AUCs were estimated.

2.6. Data Submission

Data submissions conformed to published guidelines (Martinez-Bartolome et al., 2013). LCMS data are accessible at http://proteomecentral.proteomexchange.org, identifier PXD000582. MRM data are available at http://www.peptideatlas.org/passel/, identifier PASS00343 for the MRM assay qualification data on the discovery samples, PASS00344 and PASS00959 for verification 1 and 2, respectively. Expression data are accessible at http://enews.patricbrc.org/niaid-clinical-proteomics/, identifier ZDN for the discovery data, ZHH for the MRM qualification data, and ZHS for Verification 1 data.

3. Results

3.1. Demographics and Discovery Phase

Peripheral blood specimens previously collected from 172 TB patients and household contacts were used for this study. Their clinical characteristics are shown in Table 1. The discovery phase was conducted by a cross-sectional analysis of proteomic expressions among baseline samples from 37 index cases with active TB, 8 TST-negative contacts who remained TST-negative and non-infected (NI), 21 TST-negative contacts who later converted their TST at either 3 or 6 months (converters), and 38 TST-positive contacts with LTBI at baseline. Fig. 1

	Protein	TST+ CONVERSION AT M3 M3 vs D1	M6 vs D1			Protein	TST+ CONVERSION AT M3 M3 vs D1	M6 vs D1
Acute phase	ORM1	1.66	1.23			MMP2	1.37	1.37
	CD163	1.35	1.22			ECM1	1.43	1.35
Inflammatory response	SOD3	-	-			LRG1	1.31	1.20
	CPN1	1.26	1.10			LUM	1.33	1.29
	ATRN	1.26	1.13			BTD	1.24	1.17
	S100A8	1.11	1.04			COLEC11	1.26	1.22
	S100A9	1.17	1.07		Tissue development & remodelling	COMP	1.35	1.31
Complement	C7	1.43	1.34			FCGBP	1.36	1.04
	CR2	1.19	1.09			CD248	1.39	1.43
	FCN3	1.16	0.93			CNTN1	1.18	1.27
	MBL2	1.30	1.10			COL1A1	1.27	1.30
Coagulation	VWF	1.58	1.38			MINPP1	1.27	1.20
	HABP2	1.42	1.19			PTPRG	1.24	1.17
	FLNA	1.05	0.96			MEGF8	1.26	1.14
	SERPINA5	1.28	1.18			VASN	1.21	1.20
	PROCR	1.30	1.18			IGF2	-	-
	PLEK	1.02	1.00			ENO1	1.15	1.17
Oxidative stress	GPX3	1.30	1.22		Cell proliferation	IGFBP3	1.26	1.17
	HYOU1	1.32	1.20			BCHE	1.23	1.12
	PON1	1.18	1.03			AK1	1.18	1.15
	SEPP1	1.25	1.19			ZYX	-	-
	S100A12	1.05	0.94			POSTN	1.49	1.54
Innate immunity	LBP	1.48	1.17			MYOC	1.37	1.33
	PGLYRP2	1.34	1.31			TNXB	1.25	1.27
	VNN1	1.32	1.13			TGFBI	1.31	1.19
Adaptive immunity	THBS4	1.45	1.47			ADAMDEC1	1.31	1.21
	CD14	1.28	1.21		Cell migration & adhesion	CDH13	1.31	1.27
	PROS1	1.24	1.18			IGFALS	1.35	1.29
	LCP1	1.26	1.17			OMD	1.31	1.33
	VCAM1	1.30	1.27			DSG2	1.23	1.22
	SECTM1	1.31	1.32			NID1	1.24	1.18
	LPA	1.18	1.22			GPLD1	1.28	1.11
	CPB2	1.23	1.10			MCAM	1.23	1.12
	DBH	1.28	1.11			CLSTN1	1.30	1.26
Lipid metabolism	APOA4	1.79	1.86			SELL	1.19	1.15
	APOE	1.69	1.52			CDH5	1.30	1.26
	APOA1	1.60	1.41		Glycosylation	FUCA2	1.30	1.27
	PCSK9	0.98	1.11			MAN1A1	1.27	1.16
	PRG4	1.23	1.07			PCOLCE	1.32	1.26
	APOD	1.29	1.04			CPN2	1.33	1.22
	APOC3	1.11	1.07		Proteolysis	CST6	1.25	1.21
	LCAT	1.25	1.10			CNDP1	1.26	0.96
Glycolysis	LDHA	1.33	1.27			CPQ	1.23	1.33
Transport	CLEC3B	1.30	1.28			PEPD	1.24	1.18
	PTGDS	1.24	1.25			MST1	1.25	1.15
	CACNA2D1	1.23	1.13		Other	CKM	1.42	1.64
	LAMP1	1.21	1.18			YWHAE	1.15	1.18
	HBA1	1.68	1.11					

Fig. 3. Changes in plasma protein expression after TST-positive conversion. Shown are protein expression change ratios in subjects (N = 19) with samples at baseline (D1), 3 months (M3), and 6 months (M6) that had converted to TST-positive at M3. The significant changes (p < 0.05) are color-coded red for increased and blue for decreased expression ratios, respectively.
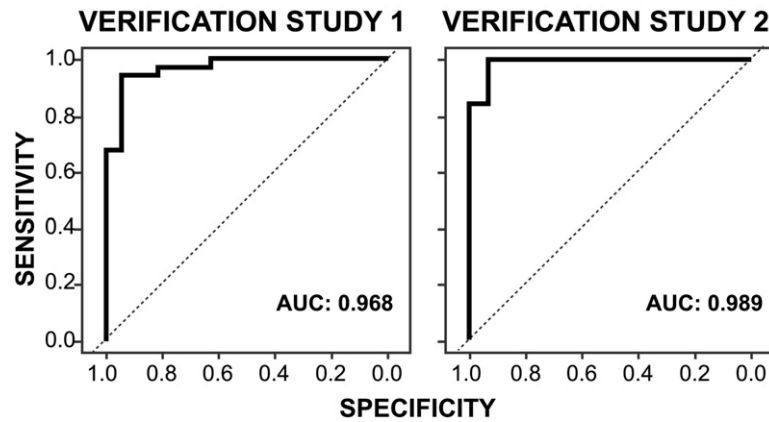
Fig. 4. Small plasma protein biomarker panels can predict TST conversion. Combinations of up to 6 proteins using the 159 targeted proteins detected were used to define biomarker panels able to predict conversion with AUC > 0.85. The initial verification study (Verification 1: $N = 52$ subjects, 3 longitudinal samples per subject) defined high performing biomarker panels whose performance was confirmed using an independent set of longitudinal samples (Verification 2: $N = 16$ subjects, 3 longitudinal samples per subject). Shown is the performance of one 6-protein panel in prediction of TST-positive conversion in both verification studies. The AUC values shown did not differ significantly from each other.

shows cross-sectional comparisons of changes in blood protein expression ratios from the baseline samples. The significant changes ($p < 0.05$) are color-coded. We found that when compared to the NI group, the number of differentially expressed proteins increased by the following group order: converters, LTBI, and active TB. The largest absolute differences in protein expression were observed between NI participants and patients with active TB. We also found differences in protein expression when looking at changes grouped by biological processes. Most protein changes associated with inflammation, immune response, tissue repair, cellular migration and proliferation were observed in subjects with active TB. Among the LTBI and converter groups, smaller changes were observed in these processes, as well as changes in proteins associated with lipid metabolism and the innate immune response (Fig. 1). The results are consistent with a low level of distinct observable changes in *Mtb* infection, including very recent infection, exhibited by the baseline TST-negative future converters who were in the process of developing LTBI.

3.2. MRM-MS Assay Development and Initial Testing (Verification 1)

A targeted MRM-MS assay was developed for 159 proteins selected from each comparison in the cross-sectional discovery phase using the following combination of biological and statistical criteria. All of the significantly differentially expressed proteins from the baseline converter vs NI and LTBI vs NI comparisons were selected along with the most differentially expressed proteins from comparisons to active TB. We also included the significant proteins described in the MRM-MS assay of Achkar et al. (2015). Re-analysis of the discovery samples with the MRM-MS assay confirmed the differential expression observed in the discovery phase (Supplemental Fig. 1). Next, an independent set of 161 longitudinal samples collected from 52 subjects who were initially all TST-negative and either converted to TST-positive ($N = 37$) or remained TST-negative ($N = 15$) were tested with the prioritized candidate biomarkers. The changes in plasma proteins at baseline and at a minimum of 2 additional time points over a 6 month period were calculated. Fig. 2 shows protein expression change ratios between each

individual's baseline TST-negative sample and the longitudinal TST-positive conversion sample or corresponding TST-negative sample. The significant changes ($p < 0.05$) are color-coded. As expected, subjects who became TST-positive had a more extensive host response than the subjects that remained TST-negative. In an analysis of subjects that converted by 3 months, most of the proteins associated with inflammatory, innate and adaptive immune responses were elevated at month 3, when TST conversion was recorded, but returned to baseline levels by month 6. In contrast, most of the proteins associated with cellular movement and tissue repair remained elevated through month 6 (Fig. 3). Converters were treated with isoniazid preventive therapy for 9 months after TST conversion, and this may have had an effect on protein expression measured at month 6.

3.3. Selection of Protein Panels Predicting TST Conversion and Validation Testing (Verification 2)

Next, the ability of plasma protein biomarker panels to predict TST conversion was evaluated. Classification analysis of combinations of all 159 targeted proteins was conducted using multivariable logistic regression with multiple training and testing subsets to classify a patient as either a TST-negative persister or a TST converter, based on the patient's baseline (TST-negative) plasma protein expression. This classification revealed a small pool of defined panels that were able to predict TST conversion with an AUC > 0.85. These high performing protein combination signatures were further tested using an independent second set of 48 longitudinal samples from 16 household contacts. Fig. 4 shows the performance in prediction of TST conversion in both verification studies of one representative high-performing panel containing proteins that were frequently observed in the high-performing panels. The composition of this panel is shown in Table 2.

4. Discussion

In this study that included analysis of serum and plasma protein expression from 172 TB patients and their household contacts, we found that the extent of host response changes was associated with progression from *Mtb* exposure to infection, with the most widespread changes occurring during active TB. Early infection was associated with specific host response processes related to inflammation, immune-response, and lipid metabolism. These specific changes allowed for the identification of protein panels predicting TST conversion. Other host response processes, such as those associated with tissue remodelling and cellular migration became most differentially expressed during active disease.

Table 2
Composition of a predictive biomarker panel.

Protein	Functional category
CLEC3B	Transport
ECM1	Tissue development & remodelling
IGFALS, SELL	Cell migration & adhesion
IGFBP3	Cell proliferation
VWF	Coagulation

These observations were consistent with previous reports (Achkar et al., 2015; De Groote et al., 2013).

Lipid metabolism was the only host process with a qualitatively different induction early in *Mtb* infection. Host lipid metabolism is known to be significantly impacted by *Mtb* infection. Fatty acids are accessed by mycobacteria during host cell infection (Lee et al., 2013), and host genes involved in lipid metabolism are highly expressed in caseous TB granulomas (Kim et al., 2010). Host lipid metabolism is also systemically affected by the acute phase response, for example by reducing levels of high density lipoprotein (HDL) and proteins such as LCAT (Khovidhunkit et al., 2004). Early in infection we observed an increase in lipid metabolism associated proteins which was reversed during active disease. This suggests that initially plasma levels of these proteins may reflect localized remodelling in the lung, an effect later overwhelmed by the influence of the acute phase response during active disease. We conclude that lipid metabolism associated peripheral blood proteins may be used to follow progression of *Mtb* infection.

Small panels of candidate biomarkers representing key host response processes were able to predict 3–6 months ahead of time the conversion from a TST-negative to a TST-positive state and thus predicted the establishment of LTBI. The key host response processes were represented by several different host proteins, though the most frequently observed are presented in the panel reported in Table 2. This panel included two proteins, CLEC3B and ECM1, involved in tissue remodelling (Holtet et al., 1997; Chan, 2004). ECM1 levels have also been linked with response to TB therapy (Nahid et al., 2014). Two other panel proteins, IGFALS and IGFBP3, are members of the insulin-like growth factor (IGF) pathway. Both proteins are found in ternary complexes with IGF, which substantially prolong IGF half-life and help regulate its activity (Boisclair et al., 2001). In active TB, we observed that IGFALS and IGFBP3 as well as IGFBP5, IGFBP6, IGF1 and IGF2 were reduced. Plasma levels of IGFBP3 have also been shown to be reduced in active *M. leprae* infection (Rodrigues et al., 2011) but increased in PBMC cultures from LTBI subjects (Stern et al., 2009). The IGFs display pleiotropic properties and can be negatively regulated by pro-inflammatory cytokines (O'Connor et al., 2008), suggesting that the elevated inflammatory response during active TB might be responsible for the observed reduction. Early in infection, however, IGF2, IGFBP3, and IGFALS were elevated in subjects that would become TST-positive, although that appeared to be a short-term change as only IGFBP3 remained significantly elevated 3 months after TST conversion. Currently it remains unclear which of the many functions of the IGF pathway are induced at this stage of *Mtb* infection. The two remaining panel proteins, SELL and VWF, have roles in the host immune response. SELL is involved in leukocyte homing and infiltration (Wedepohl et al., 2012), and was part of a host blood panel of proteins able to distinguish active TB from other respiratory diseases (Achkar et al., 2015). VWF promotes adhesion of platelets to sites of vascular injury, and is thus involved not only in tissue repair but also in the modulation of the inflammatory response (Lenting et al., 2012; Schattner, 2014). However, it is unclear if this number of host proteins represented the minimum combination necessary. Further studies may help elucidate this question.

Limitations of this study include that specimens used for the discovery and validation studies were from HIV-negative participants in a single cohort, and may not be generalizable to other settings and populations. The second verification sample size was small, which could limit the reproducibility of our findings. Additionally, household contacts were identified after the index case was diagnosed with active TB, and therefore were likely exposed and potentially already had early *Mtb* infection at the time of their baseline evaluation. Therefore, our TST-negative household contacts do not represent truly unexposed persons at baseline. Despite this we were able to identify protein signatures that predicted TST conversion, and these could be useful in TB contact investigations to identify and treat TST-negative contacts who would later develop LTBI. Similarly, given the extent of *Mtb* exposure in household contacts, as well as in the general population of this study, the NI

contacts could represent a subgroup of people who may be resistant to LTBI (Ma et al., 2014). These individuals might have different peripheral protein expression from persons who are not exposed to *Mtb*. Alternatively, some may have had depressed immunity due to conditions such as malnutrition. Given these limitations, further study and verification of these protein signature biomarkers in a low TB-incidence population would be useful. This would allow for further distinction of very recent and remote *Mtb* infection. Strengths of our study include a well-characterized epidemiological cohort of patients with newly-diagnosed culture-confirmed smear positive pulmonary TB, and contacts followed serially for TST conversion with testing done using standardized methods by a well-trained staff.

The discovery of peripheral blood protein signatures capable of predicting the development of LTBI could offer substantial clinical benefit. Use of these protein signatures in a clinical assay would enable the rapid identification of TB contacts who will develop LTBI and who would most benefit from preventive therapy, eliminating the need for follow-up testing of large numbers of TB contacts. Additionally, as LTBI represents an immunologically controlled infection, these protein biomarkers may be useful in predicting treatment response and relapse. We plan to further test these biomarkers of LTBI in low TB incidence populations and investigate their association with treatment response in active TB.

Author Contributions

Study conception and design: CMB, WHB, EP.
Acquisition of data: MN, BO, HKM, MLJ, CMS.
Analysis and interpretation of data: AMM, LLM, IR, MM, PK, SL, PT, CMB, WHB, EP.
Drafting of manuscript: CMB, EP.
Critical revision: CMB, EP, WHB, BO, CMS.

Conflicts of Interest

AM, IR, MM, PK, SL, PT, and EP were employees of Caprion Biosciences Inc. at the time the work was performed. A patent application with CMB, WHB, and EP as inventors was also submitted.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.ebiom.2017.06.019.

References

Achkar, J.M., Cortes, L., Croteau, P., Yanofsky, C., Mentinova, M., Rajotte, I., Schirm, M., Zhou, Y., Junqueira-Kipnis, A.P., Kasprowicz, V.O., Larsen, M., Allard, R., Hunter, J., Paramithiotis, E., 2015. Host protein biomarkers identify active tuberculosis in HIV uninfected and co-infected individuals. EBioMedicine 2, 1160–1168.

Boisclair, Y.R., Rhoads, R.P., Ueki, I., Wang, J., Ooi, G.T., 2001. The acid-labile subunit (ALS) of the 150 kDa IGF-binding protein complex: an important but forgotten component of the circulating IGF system. J. Endocrinol. 170, 63–70.

Chan, I., 2004. The role of extracellular matrix protein 1 in human skin. Clin. Exp. Dermatol. 29, 52–56.

Chin, D.P., Deriemer, K., Small, P.M., De Leon, A.P., Steinhart, R., Schecter, G.F., Daley, C.L., Moss, A.R., Paz, E.A., Jasmer, R.M., Agasino, C.B., Hopewell, P.C., 1998. Differences in contributing factors to tuberculosis incidence in U.S. -born and foreign-born persons. Am. J. Respir. Crit. Care Med. 158, 1797–1803.

De Groote, M.A., Nahid, P., Jarlsberg, L., Johnson, J.L., Weiner, M., Muzanyi, G., Janjic, N., Sterling, D.G., Ochsner, U.A., 2013. Elucidating novel serum biomarkers associated with pulmonary tuberculosis treatment. PLoS One 8, e61002.

Hochberg, Y., Tamhane, A.C., 2008. Multiple Comparison Procedures. Wiley.

Holtet, T.L., Graversen, J.H., Clemmensen, I., Thogersen, H.C., Etzerodt, M., 1997. Tetranectin, a trimeric plasminogen-binding C-type lectin. Protein Sci. 6, 1511–1515.

Hunter, J.M., Paramithiotis, E., 2010. Protein biomarker quantification by mass spectrometry. Expert Opin. Med. Diagn. 4, 11–20.

Keeping, E.S., 1995. Introduction to Statistical Inference. Dover Publication.

Khovidhunkit, W., Kim, M.S., Memon, R.A., Shigenaga, J.K., Moser, A.H., Feingold, K.R., Grunfeld, C., 2004. Effects of infection and inflammation on lipid and lipoprotein metabolism: mechanisms and consequences to the host. J. Lipid Res. 45, 1169–1196.

Kim, M.J., Wainwright, H.C., Locketz, M., Bekker, L.G., Walther, G.B., Dittrich, C., Visser, A., Wang, W., Hsu, F.F., Wiehart, U., Tsenova, L., Kaplan, G., Russell, D.G., 2010. Caseation of human tuberculosis granulomas correlates with elevated host lipid metabolism. EMBO Mol. Med. 2, 258–274.

Kruh-Garcia, N.A., Wolfe, L.M., Chaisson, L.H., Worodria, W.O., Nahid, P., Schorey, J.S., Davis, J.L., Dobos, K.M., 2014. Detection of Mycobacterium tuberculosis peptides in the exosomes of patients with active and latent M. tuberculosis infection using MRM-MS. PLoS One 9, e103811.

Lee, W., Vanderven, B.C., Fahey, R.J., Russell, D.G., 2013. Intracellular Mycobacterium tuberculosis exploits host-derived fatty acids to limit metabolic stress. J. Biol. Chem. 288, 6788–6800.

Lenting, P.J., Casari, C., Christophe, O.D., Denis, C.V., 2012. von Willebrand factor: the old, the new and the unknown. J. Thromb. Haemost. 10, 2428–2437.

Ma, N., Zalwango, S., Malone, L.L., Nsereko, M., Wampande, E.M., Thiel, B.A., Okware, B., Igo Jr., R.P., Joloba, M.L., Mupere, E., Mayanja-Kizza, H., Boom, W.H., Stein, C.M., Tuberculosis Research, U., 2014. Clinical and epidemiological characteristics of individuals resistant to M. tuberculosis infection in a longitudinal TB household contact study in Kampala, Uganda. BMC Infect. Dis. 14, 352.

Martinez-Bartolome, S., Deutsch, E.W., Binz, P.A., Jones, A.R., Eisenacher, M., Mayer, G., Campos, A., Canals, F., Bech-Serra, J.J., Carrascal, M., Gay, M., Paradela, A., Navajas, R., Marcilla, M., Hernaez, M.L., Gutierrez-Blazquez, M.D., Velarde, L.F., Aloria, K., Beaskoetxea, J., Medina-Aunon, J.A., Albar, J.P., 2013. Guidelines for reporting quantitative mass spectrometry based experiments in proteomics. J. Proteome 95, 84–88.

Montgomery, D.C., 2001. Design and Analysis of Experiments. Wiley.

Nahid, P., Bliven-Sizemore, E., Jarlsberg, L.G., De Groote, M.A., Johnson, J.L., Muzanyi, G., Engle, M., Weiner, M., Janjic, N., Sterling, D.G., Ochsner, U.A., 2014. Aptamer-based proteomic signature of intensive phase treatment response in pulmonary tuberculosis. Tuberculosis (Edinb) 94, 187–196.

O'Connor, J.C., McCusker, R.H., Strle, K., Johnson, R.W., Dantzer, R., Kelley, K.W., 2008. Regulation of IGF-I function by proinflammatory cytokines: at the interface of immunology and endocrinology. Cell. Immunol. 252, 91–110.

Pai, M., Schito, M., 2015. Tuberculosis diagnostics in 2015: landscape, priorities, needs, and prospects. J Infect Dis 211 (Suppl. 2), S21–S28.

Rodrigues, L.S., Hacker, M.A., Illarramendi, X., Pinheiro, M.F., Nery, J.A., Sarno, E.N., Pessolani, M.C., 2011. Circulating levels of insulin-like growth factor-I (IGF-I) correlate with disease status in leprosy. BMC Infect. Dis. 11, 339.

Sandhu, G., Battaglia, F., Ely, B.K., Athanasakis, D., Montoya, R., Valencia, T., Gilman, R.H., Evans, C.A., Friedland, J.S., Fernandez-Reyes, D., Agranoff, D.D., 2012. Discriminating active from latent tuberculosis in patients presenting to community clinics. PLoS One 7, e38080.

Schattner, M., 2014. Platelets and galectins. Ann. Transl. Med. 2, 85.

Stein, C.M., Hall, N.B., Malone, L.L., Mupere, E., 2013. The household contact study design for genetic epidemiological studies of infectious diseases. Front. Genet. 4, 61.

Stern, J.N., Keskin, D.B., Romero, V., Zuniga, J., Encinales, L., Li, C., Awad, C., Yunis, E.J., 2009. Molecular signatures distinguishing active from latent tuberculosis in peripheral blood mononuclear cells, after in vitro antigenic stimulation with purified protein derivative of tuberculin (PPD) or Candida: a preliminary report. Immunol. Res. 45, 1–12.

Storey, J.D., 2002. A direct approach to false discovery rates. J. R. Stat Soc. Ser B 64, 479–498.

Verver, S., Warren, R.M., Munch, Z., Vynnycky, E., Van Helden, P.D., Richardson, M., Van Der Spuy, G.D., Enarson, D.A., Borgdorff, M.W., Behr, M.A., Beyers, N., 2004. Transmission of tuberculosis in a high incidence urban community in South Africa. Int. J. Epidemiol. 33, 351–357.

Wedepohl, S., Beceren-Braun, F., Riese, S., Buscher, K., Enders, S., Bernhard, G., Kilian, K., Blanchard, V., Dernedde, J., Tauber, R., 2012. L-selectin–a dynamic regulator of leukocyte migration. Eur. J. Cell Biol. 91, 257–264.

Whalen, C.C., Chiunda, A., Zalwango, S., Nshuti, L., Jones-Lopez, E., Okwera, A., Hirsch, C., Peters, P., Boom, W.H., Mugerwa, R.D., 2006. Immune correlates of acute Mycobacterium tuberculosis infection in household contacts in Kampala, Uganda. Am. J. Trop. Med. Hyg. 75, 55–61.

World Health Organization, 2016. Global Tuberculosis Report 2016. World Health Organization, Geneva.

Zhang, X., Liu, F., Li, Q., Jia, H., Pan, L., Xing, A., Xu, S., Zhang, Z., 2014. A proteomics approach to the identification of plasma biomarkers for latent tuberculosis infection. Diagn. Microbiol. Infect. Dis. 79, 432–437.